

Machine Learning System on Chip (MLSoC)

Highlights

First purpose-built MLSoC™ platform with a software-centric approach providing effortless machine learning (ML) deployment for the embedded edge ML market.

Application development and evaluation:

- Supports legacy applications, as well as future ML use cases with high performance, low power, safe and secure ML inferencing.
- Supports and optimizes DNN models from a wide range of frameworks (TensorFlow, PyTorch, ONNX, MXNet etc.).
- Optimized and comprehensive ready to use Model Zoo.
- Ability to run up to four distinct DNN models concurrently.
- Optimization, visualization and debug tools.
- Software tools and libraries for easy end-to-end application development.
- PCIe development board available.

The MLSoC provides:

- Best-in-class DNN inference efficiency– 500 FPS/W ResNet-50 v1, batch size 1.
- Up to 50 TOPS for neural network computation.

Available in FCBGA 1369 balls; 31mmx31mm, 0.8mm pitch.

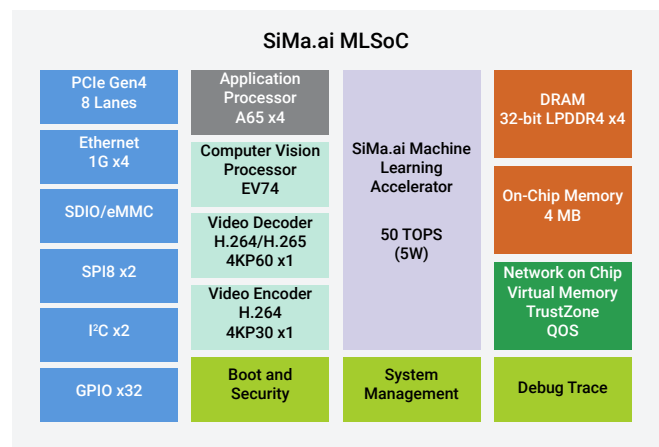
Target Applications:

- Smart Vision
- Robotics and Industry 4.0
- Drones
- Autonomous Vehicles
- Healthcare
- Government Sector

Overview

The SiMa.ai MLSoC delivers high-performance, effortless machine learning for embedded edge applications. Built on 16nm technology, the MLSoC's processing system consists of computer vision processors for image pre and post processing, coupled with dedicated machine learning acceleration and high-performance application processors. Surrounding the real-time intelligent video processing are memory interfaces, communication interfaces, and system management all connected via a network on chip (NoC).

The MLSoC features low operating power and high ML processing capacity, making it ideal as a standalone edge-based system controller or to add a machine learning offload accelerator for processors, ASICs and other devices.



MLSoC Architecture and Features

The MLSoC contains the following major functions:

- Machine learning accelerator (MLA) – providing 50 teraoperations per second (50 TOPS) for neural network computation at 10 TOPS/W.

- Application processing unit (APU) – a cluster of four Arm Cortex-A65 dual threaded processors operating at 1.15 GHz to deliver up to 15k Dhrystone MIPS.
- Video encoder/decoder – supports the H.264 compression standards (HVEC for decode) with support for baseline/main/high profiles, 4:2:0 pixels and 8-bit precision. The encoder supports rates up to 4KP30, while the decoder supports up to 4KP60.
- Computer vision unit (CVU) – consists of a four-core Synopsys ARC EV74 video processor supporting up to 600 16-bit GOPS.
- High-speed I/O subsystem – provides four 1 Gigabit Ethernet ports plus a PCIe Gen4 8-lane interface as an endpoint.
- Low-speed I/O subsystem – lower bandwidth interfaces such as SPI, I2C, GPIO, and SDIO/eMMC flash interfaces.
- DRAM interface system (DIS) – supporting four 32-bit LPDDR4 memory controllers and PHY interfaces operating at up to 3733 MTps.
- Boot and security unit (BSU) – provides secure key storage in an OTP memory (efuse) and key management. Supports decryption and authentication of the boot image as well as providing a security API to the user code.

Software-First Development Environment

Compiling an ML-trained model to target particular hardware can be challenging if the software toolchain and hardware are not co-designed from the very beginning. Our software-first approach includes carefully defined intermediate representations (including the TVM Relay IR), along with novel compiler optimization techniques. This software architecture enables SiMa.ai to support a wide range of frameworks (e.g., TensorFlow, PyTorch, ONNX, etc.) and compile over 120+ networks, thus providing customers with an effortless experience and world-class performance-per-watt results.

About SiMa.ai

SiMa.ai is a machine learning company delivering the industry's first software-centric purpose-built MLSoC platform. With push-button performance, we enable effortless ML deployment and scaling at the embedded edge by allowing customers to address any computer vision problem while achieving 10x better performance at the lowest power. Initially focused on computer vision applications, SiMa.ai is led by technologists and business veterans backed by a set of top investors committed to helping customers bring ML on their platforms. For more information, visit www.SiMa.ai.



SiMa Technologies, Inc.
226 Airport Parkway, Suite 550
San Jose, CA 95110
mlsoc@sima.ai

SiMa.ai India Private Limited
Bagmane Tech Park Unit 02
2nd Floor, B Wing, Laurel Building
C V Raman Nagar, Bengaluru, Karnataka – 560093